TUPM 6.7

AN INVESTIGATION OF AN MPEG-4 EMBEDDED SPATIAL DIGITAL WATERMARK DEVELOPED FOR TAMPER DETECTION AND CHARACTERIZATION

D.A. Winne, H.D. Knowles, D.R. Bull and C.N. Canagarajah

University of Bristol, Centre for Communications Research, Merchant Venturers Building, Woodland Road, Bristol, BS8 1UB, UK

ABSTRACT

The widespread adoption of digital video techniques has generated a requirement for authenticity verification in applications such as criminal evidence, insurance claims and commercial databases. This paper investigates the performance, effectivity and efficiency of a spatial digital watermark designed for video authentication. The embedded watermark is estimated from the possible tampered frame with the use of a temporal filter and the motion information. The imperceptible digital stamp is robustly embedded prior to encoding (compression) and allows tamper characterisation of time-base attacks. The functionality of this system within an MPEG-4 implementation is evaluated.

INTRODUCTION

The vision of cameras watching your every move is close to becoming a reality, with analysts predicting a tenfold increase in CCTV in UK in the next five years [1]. Spy cameras [2] designed to target school vandals or catch burglars [3] have become a reality. Authentication has always been an important issue throughout history [4]. Now, in the world of digital imagery and video, the aid of digital watermarks has made it more accessible.

Fragile watermarks designed to detect and locate subtle modification in images have proven useful for authentication (see [5] and [6]), however they lack the ability to provide further information necessary to characterize the attack. This is especially valid for attacks on video signals. Attacks on video signals fall into two categories, [7]: The first type covers attacks that tamper with the intensity patterns of the video, e.g. compression, noise, etc. The second type, time-base tampering, disrupts the frame sequencing, e.g. frame cuts, swapping, deletion or foreign frame insertion.

It is not straightforward to apply an image watermarking scheme on a video system as a video system is more than an ordered collection of images. Hybrid video compression systems encode each frame of a sequence differentially. There are temporal and stop frame visibility issues. There is a possibility of reference frame information leakage and often real-time application constraints apply.

VIDEO WATERMARKING SYSTEM

The system designed to detect and characterise time-base attacks was initially described in [8], further refined in [10], and is embedded in a standard MPEG-4 codec. The watermark, shaped by a well-known texture masking function [9], is embedded by signal adaptive addition in the spatial domain, avoiding the need for a drift compensation system:

$$y = x + \alpha n \tag{1}$$

where α represents the overall strength factor and is linked to the peak signal to noise ratio between the previous encoded frame and its original. This adapts the watermark strength to bit-rate variations. An estimate of the embedded watermark is generated as follows:

$$\tilde{n} = \hat{y} - \tilde{x} \tag{2}$$

where $\tilde{x}$ is a estimate of the original frame x and is a de-noised representation of the watermarked and possibly tampered frame $\hat{y}$.

The reference frame watermark leakage distortion is circumvented by embedding the same static watermark in low motion (often background) area of subsequent P-frames and by adapting a dynamic watermark per frame in the active regions. The mean absolute difference between two adjacent frame in combination with a threshold differentiates between the static and dynamic area.

The dynamic watermark extraction procedure combines a temporal filter with the motion information. The estimate of the static watermark from each individual frame is merged together to form an improved estimate. The similarity between the extracted and embedded watermark is measured with normalised correlation.

The watermark payload depends on the activity and bit-rate of the sequence. The extracted information can be utilized to differentiate between rate-controlled or illicit frame skipping. The static watermark allows detection of GOP deletion and contains global information such as bit-rate, number of skipped frames in previous GOP or time information.

0-7803-7721-4/03 $17.00 © 2003 IEEE

The dynamic watermark maintains a unique frame based watermark and allows detection of inter GOP tampering such as frame removal. The extracted information forms a measure for the authenticity of the sequence.

EVALUATION

The Receiver Operating Characteristic (ROC) allows us to compare the performance of different watermarking systems. It displays the probability of false positive, F_p, in function of the probability of false negative, F_n. One estimates F_p and F_n by calculating the surface under the model distribution with the boundaries set to *]-inf, T]*, where T represents the detection threshold. F_p is estimated by assuming that the watermark is drawn from a radially symmetric distribution [11]. F_n is estimated by assuming that the normalised correlation between the extracted watermark and the embedded watermark can be modelled parametrically as a bimodal Gaussian Mixture model. The ideal ROC should lie as close as possible to the axes, with F_p=F_n=0.

Figure 1 displays the ROC of the dynamic watermark of the *Silent* test sequence. The left solid black graph indicates the trade-off between F_p and F_n when the watermarked sequence is compressed at *1Mbps*. The 4 graphs on the right illustrated the performance decline when the watermarked and encoded sequence is subsequently re-encoded (transcoded) at a bit-rate of *1Mbps, 512kbps, 256kbps* and *128kbps*. Even though the sequence was encoded again at *1Mbps*, the non-linear performance of the quantizer and the different quantization step-sizes increases the frame distortion. This influences the estimate of the embedded watermark, which results in an increase of F_n for a given F_p. Transcoding at a lower bit-rate decreases the performance rapidly.

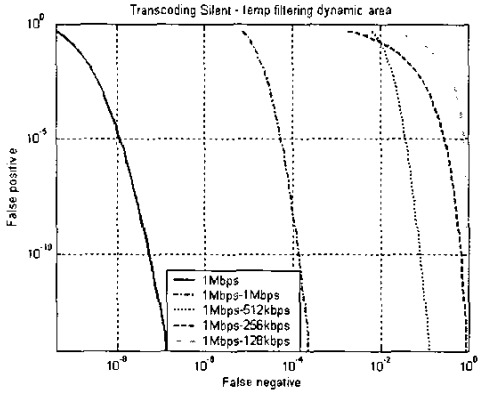

Figure 1: ROC of Silent test sequence, W extraction with temporal filtering Watermarked sequence encoded at 1Mbps and subsequently re- encoded at 1Mbps, 512kbps, 256kbps and 128kbps

CONCLUSIONS

This paper describes in depth the technical design, implementation issues and capabilities of a video watermarking system. The system reliably extracts information embedded in the spatial domain and utilises this to measure the authenticity and characterise tampering scenarios. The performance is examined in an MPEG-4 environment with a Receiver Operating Characteristic. The efficiency of the system is compared with other reported video authentication systems. The effectivity is evaluated with the use of different time-base tampering scenarios.

REFERENCES

[1] Watching your every move, Thursday, 7 February 2002, BBC News Online, http://news.bbc.co.uk/1/hi/sci/tech/1789157.stm

[2] Spy cameras target school vandals, Wednesday 5 June 2002 BBC News, Online http://news.bbc.co.uk/1/hi/sci/tech/2016772.stm

[3] Spy cameras catch burglars, Tuesday 17 December 2002, BBC News Online, http://news.bbc.co.uk/1/hi/england/2582495.stm

[4] F. Hartung and M. Kutter, "Multimedia watermarking techniques," *IEEE Proceedings*, Vol. 87, No. 7, pp. 1076-1107 July 1999.

[5] D. A. Winne, H. D. Knowles, D. R. Bull, C. N. Canagarajah, "Compression Compatible Digital Watermark algorithm for Authenticity Verification and Localization", *SPIE conference: Security and Watermarking of Multimedia Contents IV*, 2001, San Jose, paper-ID: 4675-38

[6] D. A. Winne, H. D. Knowles, D. R. Bull, C. N. Canagarajah, "Digital Watermarking in Wavelet Domain with Predistortion for Authenticity Verification and Localization", *SPIE conference: Security and Watermarking of Multimedia Contents IV*, 2001, San Jose, paper-ID: 4675-39

[7] B. C. Mobasseri, M. J. Sieffert and R. J. Simard, "Content authentication and tamper detection in digital video," *IEEE Int'l Conf. on Image Proc.*, Vancouver, paper ID:1973, Sept. 2000.

[8] Dominique A. Winne, Henry D. Knowles, David R. Bull, C. Nishan Canagarajah, "Spatial digital watermark for MPEG-2 video authentication and tamper detection", *IEEE International Conference on Acoustics, Speech and Signal Processing*, Orlando-Florida, pp 3457-3460, May 13-17.

[9] S.Voloshynovskiy, A.Herrigel, N.Baumgärtner and T.Pun, "A stochastic approach to content adaptive digital image watermarking", In International Workshop on Information Hiding, Dresden, Germany, , Lecture Notes in Computer Science, Ed. Andreas Pfitzmann, pp. 211-236, 29 September - 1 October, 1999.

[10] Dominique A. Winne, Henry D. Knowles, David R. Bull, C. Nishan Canagarajah, "An Efficient MPEG-4 video authentication scheme with temporal filtering", submitted for International Conference on Signal Processing, 2003 Barcelona

[11] M. L. Miller and J. A. Bloom, "Computing the Probability of False Watermark Detection", Proceedings of the Third International Workshop in Information Hiding, pp. 146-158, 1999